

Final Project Proposal: Empirical Banzhaf Indices

Mark Schulist

1) Inspiration

Over the past few elections, I have been disappointed in how few Americans vote. Most of them just seem to not care because their vote *doesn't matter* (of course, this would be solved with a popular vote, but oh well *democracy*). I would like to be able to respond with “well actually, your vote matters x amount relative to other people,” but I’ve never been able to find a good way of phrasing my thoughts into a convincing argument.

When we talked about power indices in class, my immediate thought was to apply them to the US states and the electoral college. Upon learning about the time complexity of Banzhaf ($\mathcal{O}(2^n)$), I became disappointed. Then, I learned about generating functions and was amazed. Of course, the paper talking about generating functions had already applied their method to the US electoral college, rendering my ideas useless...

But then I became to ponder what the power indices really *measured*. Do they measure true power, or just power in an ideal world of equally likely outcomes? Does California really have more power than Pennsylvania? In our ideal world, yes, but there’s a reason that the *AP* calls the CA race immediately but can take days to call the PA race.

2) Execution

My goal is to (using [1]), determine the **empirical Banzhaf Indices** for each state in the US. Of course, the problem of the large sample space reappears (unless I discover a nice generating function, which I doubt I will).

To get over the large sample space problem, I will divide the US into “blocks” that represent psuedo-states that have always voted together. For example, New York and California have voted the same in every election since 1988, so they can (along with many other similar states) be grouped together.

3) Bayesian Statistics

We let $i_j \in \{0, 1\}$ represent that voter j votes yes/no and $p_{i_1 \dots i_m}$ be the probability in a random election.

There are m voting members (so in the case of the US without blocks, $m = 51$).

We let $n = \sum x_{i_1 \dots i_m}$ be the total number of past elections where all m parties were present.

After much multiplication, we get the estimates for our voting probabilities.

$$\hat{p}_{i_1 \dots i_m} = \frac{1}{2} \frac{x_{i_1 \dots i_m} + x_{i'_1 \dots i'_m} + 1}{n + 2^{m-1}} \quad (1)$$

$x_{i'_1 \dots i'_m}$ represents the complement to $x_{i_1 \dots i_m}$, in other words, the number of times the particular winning coalition voted no (0).

For example, if we have $m = 2, n = 5$, we could have $x_{11} = 3, x_{00} = 1, x_{01} = 1, x_{10} = 0$. As we can see, the number of elements of x grows as 2^m , so this can get costly to compute for large values of m .

Computing $\hat{p}_{00} = \hat{p}_{11}$:

$$\frac{1}{2} \frac{3 + 1 + 1}{5 + 2^{2-1}} = \frac{5}{14} \quad (2)$$

Bibliography

[1] A. Heard and T. Swartz, “Empirical Banzhaf Indices,” *Springer Nature*, vol. 97, no. 4, pp. 701–707, 1998.